

DOCUMENT RESUME

ED 146 013

SE 022 857

AUTHOR Weaver, J. F.
TITLE Consistency of Fifth-grade Students' Solution Methods for Certain Open Distributive Sentences: Hobbs' Data Revisited. Project Paper 76-5.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO WRDCCL-PP-76-5
PUB DATE Dec 76
CONTRACT NE-C-00-3-0065
NOTE 18p.; Report from the Project on Analysis of Mathematics Instruction

EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.
DESCRIPTORS Achievement; Elementary Education; *Elementary School Mathematics; Grade 5; Mathematics Education; *Multiplication; *Number Concepts; *Research
IDENTIFIERS *Mathematical Sentences; *Properties (Mathematics)

ABSTRACT

Studies on distributive sentences by Weaver and by Hobbs are reviewed, with procedures and findings summarized. Data from Hobbs were reanalyzed to answer the question, to what extent did individual students tend to use the same solution method for all four types of open sentences? A one-way ANOVA indicated that there appeared to be no differential consistency effect associated with the context factor. However, the magnitude of the context means suggested that students exhibited a marked tendency to use just one solution method across sentences, although the method was not necessarily the same for all students. A technique using "consistency scores" affirmed these findings. The degree and type of consistency found are discussed. (MS)

* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *

ED146013

Project Paper 76-5

Consistency of Fifth-grade Students' Solution Methods
for Certain Open Distributive Sentences:
Hobbs' Data Revisited

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. F. Weaver

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

by

J. F. Weaver

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY.

Report from the Project on
ANALYSIS OF MATHEMATICS INSTRUCTION

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

December 1976

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. NE-C-00-3-0065

SE 022 857

Preface

There have been various facets to my long-standing interest in school students' solutions to open number sentences. One of these pertains to methods they use to solve such sentences, including the phenomenon of consistency or inconsistency in their choice of solution procedures.

Appendix B of a recent research report by Hobbs (1976) provides raw data relevant to this phenomenon. I take full personal responsibility for the subsequent analyses I have applied to Hobbs' data,--analyses that were not within the domain of the investigation he conceptualized.

Consistency of solution procedures may or may not be a virtue. Regardless, it is a phenomenon that merits investigation.

An HP-97 programmable printing electronic calculator was used for all computations. Two programs in particular were run: the Analysis of Variance (One-Way) program ST1-06A from HP-67/HP-97 Stat Pac I (1976), and the repeated measures program developed by Weaver (1976).

jfw

Background

A normative investigation reported by Weaver (1973), involving more than 2000 fourth-, fifth-, sixth- and seventh-grade pupils drawn from 100+ classes in 12 schools from two districts, focused on performance with open sentences of the eight types identified in Table 1. Data were derived from administering to each pupil one of the eight nine-item tests characterized in Table 2, where the first eight items involved the item types included in Table 1.

Tests were scored in terms of criterion responses (reflecting an application of distributivity) and in terms of the broader category of mathematically correct responses as defined in Figure 1. Attention also was given to categorizing responses that were not mathematically correct.

At each grade it was observed that:

1. Performance in terms of both criterion responses and mathematically correct responses was at a markedly low level.

2. Performance in terms of both criterion responses and mathematically correct responses clearly was higher for regrouping-sets items than for multiplication-addition items.

3. Performance in terms of both criterion responses and mathematically correct responses clearly was higher for right-distributive items than for left-distributive items.

4. Performance in terms of both criterion responses and mathematically correct responses was at essentially the same level for items presented horizontally and for items presented vertically.

5. An interaction involving (2) and (3) above was in evidence. The higher performance level on right-distributive items than on left-distributive items was marked within the regrouping-sets context but no pronounced tendency for this same phenomenon was observed within the multiplication-addition context.

6. The majority of mathematically incorrect responses fell within two categories:

- 6.1. A pupil response (a,b) was of the form $\underline{a} = \underline{kp}$, $\underline{b} = \underline{kq}$; or, $\underline{a} = \underline{kq}$, $\underline{b} = \underline{kp}$. This will be identified as a "category 3" response.

TABLE 1

Example Types Determined by Completely Crossing Factors A, B, C (Levels 1 and 2)

A. Mathematical context	B. Distributive form	C. Example format	
		1. Horizontal	2. Vertical
 1. Regrouping sets	1. Left	k sets of p and k sets of q are ___ sets of ___	$\begin{array}{r} k \text{ sets of } p \\ \text{and } k \text{ sets of } q \\ \hline \text{___ sets of ___} \end{array}$
	2. Right	p sets of k and q sets of k are ___ sets of ___	$\begin{array}{r} p \text{ sets of } k \\ \text{and } q \text{ sets of } k \\ \hline \text{___ sets of ___} \end{array}$
2. Multiplication-addition	1. Left	$(k \times p) + (k \times q) = \_ \times \_$	$\begin{array}{r} (k \times p) \\ + (k \times q) \\ \hline \_ \times \_ \end{array}$
	2. Right	$(p \times k) + (q \times k) = \_ \times \_$	$\begin{array}{r} (p \times k) \\ + (q \times k) \\ \hline \_ \times \_ \end{array}$

TABLE 2
Characteristics of Test Items

Test	Item #1	Item #2	Item #3	Item #4	Item #5	Item #6	Item #7	Item #8	Item #9
01	Regrouping sets				Multiplication-addition				
	RV	LH	LV	RH	RV	RH	LH	LV	AH
	RH	RV	LH	LV	LH	LV	RV	RH	AV
	LV	RH	RV	LH	RH	LH	LV	RV	AH
07	LH	LV	RH	RV	LV	RV	RH	LH	AV
	Multiplication-addition				Regrouping sets				
	RH	LV	LH	RV	LV	RH	LH	RV	AH
	LH	RV	RH	LV	RH	LV	RV	LH	AV
06	RV	RH	LV	LH	RV	LH	RH	LV	AH
	LV	LH	RV	RH	LH	RV	LV	RH	AV
	RH	RV	LH	LV	RH	LV	RV	LH	AV
	LV	LH	RV	RH	LH	RV	LV	RH	AV

Note.--L = Left-distributive form

R = Right-distributive form

A = Ambiguous distributive form (i.e., when $p = q$ in Table 1 types)

H = Horizontal example format

V = Vertical example format

The following additional condition prevailed, where k , p , and q are to be interpreted as in Table 1:

For items 1, 2, 5, and 6 of each test, $4 < k < 10$; $1 < p < 10$; $1 < q < 10$; and $p + q > 10$.

For items 3, 4, 7, and 8 of each test, $11 < k < 17$; $1 < p < 10$; $1 < q < 10$; and $p + q > 10$.

For item 9 of each test, $4 < i < 10$, $11 < j < 17$; or $11 < i < 17$, $4 < j < 10$.

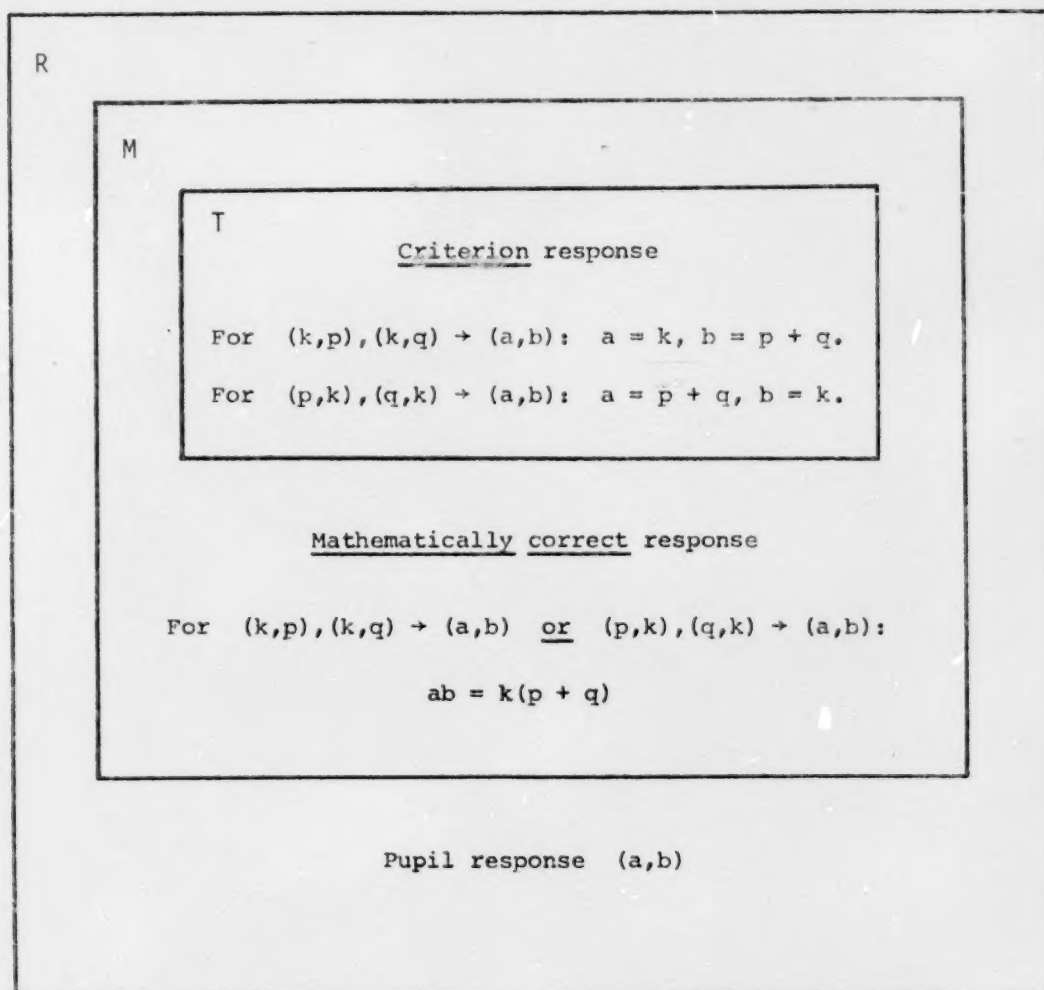


Figure 1. Defining conditions for criteria and mathematically correct responses

Note. The ordered pair (a,b) for a pupil response signifies the responses for the first (a,_) and second (_,b) blanks, respectively, associated with an open sentence.

Category 3 responses occurred more frequently within the multiplication-addition context than within the regrouping-sets context and occurred more frequently for horizontal-format items than for vertical-format items.

6.2. A pupil response (a,b) was of the form $\underline{a} = \underline{2k}$, $\underline{b} = \underline{p + q}$; or $\underline{a} = \underline{p + q}$, $\underline{b} = \underline{2k}$. This will be identified as a "category 4" response.

Category 4 responses occurred more frequently with the regrouping-sets context than within the multiplication-addition context and occurred more frequently for vertical-format items than for horizontal-format items.

The principal findings just summarized prompted Hobbs (1976) to conduct more of an in-depth investigation with a replication of some variables from the Weaver study and an extension or modification of others.

The Hobbs Investigation

Hobbs (1976) administered individually to each of 80 fifth-grade students (drawn from three schools within the same district) four open sentences in horizontal format in which a random order of item presentation was selected from the matrix represented by Table 3.

TABLE 3
Item Factors and Levels

		Distributive form	
		Left: $(k \times p) + (k \times q) = \underline{\quad} \times \underline{\quad}$	Right: $(p \times k) + (q \times k) = \underline{\quad} \times \underline{\quad}$
Number size	Small: $1 < k < 10$, $1 < p < 10$, $1 < q < 10$, $(p+q) > 10$.		
	Large: $11 < k < 20$, $11 < p < 20$, $11 < q < 20$, $(p+q) > 30$.		

Students also were assigned randomly to one of four contexts in which were presented the open sentences to be solved ($n = 20$ students/context):

- (1) Symbolic (S). Each open sentence was presented simply as, e.g.,
 $(4 \times 5) + (7 \times 5) = (\_\times\_\)$, or (OR in left distributive form)
 $(19 \times 18) + (12 \times 18) = (\_\times\_\)$.
- (2) Symbolic/Pictorial (SP). Each open sentence was presented symbolically as in (1) and in explicit association with a pictured rectangular array of small objects (tacks).
- (3) Symbolic/Verbal (SV). Each open sentence was presented symbolically as in (1) along with an explicit verbal suggestion that the distributive or "common factor" property might be used.
- (4) Symbolic/Pictorial/Verbal (SPV). Each open sentence was presented symbolically as in (1), in explicit association with a pictured array of tacks and an explicit verbal suggestion that the distributive or "common factor" property might be used.

Hobbs used the same eight categories as did Weaver (1973) to classify the procedures used by students when solving the open distributive sentences. Hobbs' findings that pertain particularly to the analyses reported in this Project Paper are summarized in Table 4 in which the number of classification categories for solution method has been abridged to four.

Three things, at least, are very apparent from Table 4:

1. Mathematically valid solution methods were attempted in less than 25% of the instances, with cell "S/P/V, Method 1" being empty.
2. More than 60% of the attempted solutions involved the same invalid procedure: method 3.
3. The relative frequency of particular solution methods was more variable across context levels than across distributive-form levels or number-size levels.

Table 4 data, however, in no way provide an answer to the question, To what extent did individual students tend to use the same solution method for all four open sentences? In view of the first part of Table 4, it also might be asked, Is the tendency (whatever it may be) to use the same solution method across open-sentence types the same regardless of the context in which the sentences were presented?

One simplistic approach to answering these questions is through a one-way ANOVA in which each data point is the number of different methods used by a particular student to solve the four open sentences presented to him/

TABLE 4
Frequency of Use of Solution Methods

Factor and level	Method				Sum
	1	2	3	Other	
Context					
Symbolic	16	12	35	17	80
Symbolic/Pictorial	15	10	47	8	80
Symbolic/Verbal	15	4	54	7	80
Symbolic/Pictorial/Verbal	0	4	62	14	80
Sum	46	30	198	46	320
Distributive form					
Left	22	15	101	22	160
Right	24	15	97	24	160
Sum	46	30	198	46	320
Number size					
Small	23	16	104	17	160
Large	23	14	94	29	160
Sum	46	30	198	46	320

Notes.

Each cell entry is the number of times a solution method was used.

Method 1: application of distributivity.

Method 2: application of a mathematically valid procedure other than distributivity.

Method 3: application of the invalid procedure whereby an (a,b) response is intended as $a = kp$ and $b = kq$.

Other: Any method distinct from 1 or 2 or 3.

her within a particular context. Such "scores" range from a high of 1 (all methods the same) to a low of 4 (all methods different). Context means and the associated ANOVA are summarized in Table 5. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ is

tested against H_1 : not H_0 at $\alpha = .05$, for which $F_{.95}(3,76) = 2.72$.

TABLE 5
Analysis for Consistency Data I

Context	Mean	Variance estimate	Source of variation	SS	df	MS	F
S	1.25	.188	Between contexts	.54	3	.18	.69
SP	1.25	.288	Within contexts (error)	19.65	76	.26	
SV	1.20	.460					
SPV	1.05	.047	Total	20.19	79		
Grand mean	1.19						

Clearly, H_0 is not rejected in favor of H_1 . There appears to be no differential consistency effect associated with the context factor.

The magnitude of the context means--close to 1.00--suggests that students exhibited a rather marked tendency to use just one solution method for all four open sentences, although the particular method used was not necessarily the same for all students. Further evidence of this degree of consistency is seen in the fact that across contexts 85% of the students used some one procedure to solve each of the four open sentences (75% among S-context students, 80% among SP-context students, 90% among SV-context students, and 95% among SPV-context students).

A second approach. Consider the following set of combinations:

a,a,a,a
a,a,a,b
a,a,b,b
a,a,b,c
a,b,c,d

where a, b, c, d represent distinct solution methods. It is true that the preceding approach to consistency does not distinguish between the second and third rows of combinations just listed. In each of these instances two distinct solution methods are evident, but it can be argued that the second row reflects greater consistency of methods used than does the third row.

Based upon that argument, it was decided to assign to each pupil a "consistency score" based upon the preceding set of solution-method combinations (regardless of the particular methods involved or the order in which used):

Score 4 if a,a,a,a

Score 3 if a,a,a,b

Score 2 if a,a,b,b

Score 1 if a,a,b,c

Score 0 if a,b,c,d.

Table 6 summarizes an analysis based upon student "consistency scores" as just defined.

TABLE 6
Analysis for Consistency Data II

Context	Mean	Variance estimate	Source of variation	SS	df	MS	F
S	3.60	.540	Between contexts	.95	3	.32	.56
SP	3.70	.510	Within contexts	43.00	76	.57	
SV	3.70	.910	(error				
SPV	3.90	.190	Total	43.95	79		
Grand mean	3.73						

As before, there clearly is no cause to reject H_0 in favor of H_1 . There appears to be no differential consistency effect associated with the context factor.

Again, the magnitude of the context means--close to 4.00--suggests that students tended to a rather marked degree to use one particular solution method (not necessarily the same one for each student, however) across the four open sentences.

The two different approaches to assessing consistency lead essentially to the same conclusion,--possibly in large measure because there were very few instances in which the distinction between "consistency scores" of 3 and 2 needed to be made in the raw data.

Consistency among items. Thus far consideration has been given to consistency of solution method with respect to open-sentence context (without regard for item type). Now attention is directed to consistency of solution method with respect to item type.

Table 4 indicated 46 uses of Method 1, 30 uses of Method 2, and 198 uses of Method 3. The 46 uses of Method 1 were all among only 13 students; the 30 uses of Method 2, among only eight students; and the 198 uses of Method 3, among 54 students. These data in themselves suggest a high degree of consistency in students' use of Method 1 or Method 2 or Method 3 when solving the four open sentences.

However, it nevertheless might be that--especially in connection with Method 3, and possibly Method 1--one or another of the item types prompted less use of a particular method than did the other item types. This possibility was examined in the following way:

Consider first the 13 students each of whom used Method 1 at least once when solving the four open sentences. For each of these students, if Method 1 was used to solve a particular open-sentence type, that item was "scored" 1; if not, it was "scored" 0. Thus, there was generated for each of the 13 students an ordered set of four dichotomous scores based on use of Method 1.

Similar procedures were followed for the eight students each of whom used Method 2 at least once when solving the four open sentences, and also for the 54 students each of whom used Method 3 at least once when solving the four open sentences.

Context was disregarded when generating each of the preceding three classes of ordered sets of scores. Table 7 summarizes consistency data based upon open-sentence type means for each of the three methods. We may designate these means as $\bar{X}_{11}$, $\bar{X}_{12}$, $\bar{X}_{13}$, $\bar{X}_{14}$, $\bar{X}_{21}$, $\bar{X}_{22}$, . . . , $\bar{X}_{33}$, $\bar{X}_{34}$ reading from left to right across rows. If desired, each mean may be multiplied by 100 and interpreted as a percent.

The calculator program used for the repeated measures design involved in Table 7 (Weaver, 1976) generates both the F statistic and the Cochran Q statistic, so each is reported for informational purposes. The complete ANOVA data associated with the calculation of the Table 7 F's and Q's are given in Appendix A.

TABLE 7
Consistency Data for Item Types

Method	n	Mean				F	Q
		Left distributive form		Right distributive form			
		Small number size	Large number size	Small number size	Large number size		
1 ^a	13	.846	.846	.923	.923	.38	1.20
2 ^b	8	1.000	.875	1.000	.875	1.00	3.00
3 ^c	54	.981	.889	.944	.852	3.81	10.88

$$^a F_{.95}(3,36) = 2.86 \quad \text{and} \quad \chi^2_{.95}(3) = 7.82$$

$$^b F_{.95}(3,21) = 3.07 \quad \text{and} \quad \chi^2_{.95}(3) = 7.82$$

$$^c F_{.95}(3,159) = 2.67 \quad \text{and} \quad \chi^2_{.95}(3) = 7.82$$

For Method 1 there appears to be no cause to reject $H_0: \mu_{11} = \mu_{12} = \mu_{13} = \mu_{14}$ in favor of $H_1: \text{not } H_0$; likewise for Method 2 there appears to be no cause to reject $H_0: \mu_{21} = \mu_{22} = \mu_{23} = \mu_{24}$ in favor of $H_1: \text{not } H_0$ ($\alpha = .05$). For Method 3, however, at $\alpha = .05$ there clearly is cause to reject $H_0: \mu_{31} = \mu_{32} = \mu_{33} = \mu_{34}$ in favor of $H_1: \text{not } H_0$.

The preceding conclusions are not unrelated to the following observations regarding the raw data: Of the 13 students who used Method 1 at least once, 10 of them used that method for all four open sentences; of the eight students who used Method 2 at least once, seven of them used that method for all four open sentences; and of the 54 students who used Method 3 at least once, 45 of them used that method for all four open sentences. Student performance involving Method 2 provided virtually no opportunity for variation among column means, but this was not the case for Methods 1 and 3.

In Conclusion

The relatively high degree of consistency observed in the methods used by individual students to solve the open distributive sentences may or may not be a virtue.

In the case of Method 1, consistency of use reflects a desirable recognition and application of a mathematical pattern. That is good.

In the case of Method 2, consistency of use reflects at least a valid interpretation of the open sentences and a valid approach to their solution. That, too, is good even though recognition and application of the distributivity pattern might be considered more desirable. From an instructional standpoint I hypothesize that consistency with Method 2 is more facilitating than inhibiting in attempts to develop with such students the validity and application of Method 1.

In the case of Method 3, is consistency of use a virtue? Certainly it is not "virtuous" for students to consistently do the wrong thing! But from an instructional point of view, is consistency with Method 3 more facilitating or more inhibiting than inconsistency with Method 3 would be when attempting to develop with such students the validity and application of Method 1? An established routine often is considered to be more resistant to change than one that is not well established.

Hobbs (1976) indicated that some students who used Method 3 in their attempts to solve open number sentences did so for items within the small-number domain but not for items within the large-number domain. "The change in solution procedures was due primarily to a lack of students' ability to multiply large numbers (two-digit numbers)." (p. 107) What implication, if any, has this for instruction?

Remedial instruction pertaining to the distributive property of multiplication over addition will not be the same for all students who are in need of such remediation. One differentiating factor to take into consideration may be the degree of consistency in students' use of particular invalid procedures such as Method 3. Explorations are needed to ascertain whether consistency with an invalid procedure is a facilitating or inhibiting factor in the remedial situation.

References

- Hobbs, C. E. An analysis of fifth-grade students' performance when solving selected open distributive sentences. Technical Report No. 397. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.
- HP-67/HP-97 Stat Pac I. Cupertino, CA: Hewlett-Packard, 1976.
- Weaver, J. F. Pupil performance on examples involving selected variations of the distributive idea. Arithmetic Teacher, 1973, 20, 697-704.
- Weaver, J. F. An HP-97/67 program for analyses of data from a single-factor design with repeated measures. Project Paper 76-4. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.
- Winer, B. J. Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill, 1971.

Appendix A

ANOVA Data for Table 7

		<u>Method 1</u>	<u>Method 2</u>	<u>Method 3</u>
Between students:	SS	2.500 ***	0.875 ***	8.500 ***
	df	12.000 ***	7.000 ***	53.000 ***
	MS	0.208 ***	0.125 ***	0.160 ***
Within students:	SS	2.500 ***	1.000 ***	8.000 ***
	df	39.000 ***	24.000 ***	162.000 ***
	MS	0.064 ***	0.042 ***	0.049 ***
Between items:	SS	0.077 ***	0.125 ***	0.527 ***
	df	3.000 ***	3.000 ***	3.000 ***
	MS	0.026 ***	0.042 ***	0.179 ***
Residual:	SS	2.423 ***	0.875 ***	7.463 ***
	df	36.000 ***	21.000 ***	159.000 ***
	MS	0.067 ***	0.042 ***	0.047 ***
Total:	SS	5.303 ***	1.875 ***	16.500 ***
	df	51.000 ***	31.000 ***	215.000 ***
	F	0.761 ***	1.000 ***	3.614 ***
	Q	1.200 ***	3.000 ***	10.875 ***